				Same text as #212. • 3101-3102: It is correct that ever/never considerations does not allow a dose response analysis. It is not correct, however, that it is necessarily linked with high misclassification rates. If fact, if the issue is to group individuals into those who may have ever been exposed and those who have not, it may do that quite well. In any case, you cannot determine the misclassification rate for such a grouping simply because there are only two categories. *EFSA Response:* *One problem often associated with this design is the lack of information about other exposures that may have affected the outcomes and the absence of a control population of subjects known not to have been exposed to any other toxins that might have similar effects. Although not a human study such a problem has made interpretation of a recent field studies of neonicotinoid pesticides on bees in 3 countries where measurement of the effects of one active substance became uncertain when evidence of a different neonicotinoid in the hives was discovered alongside the neonicotinoid of interest to the researchers.* • 3102-3104: Case-control studies and misclassification. While it is true that bias in case-control studies can be differential, it can be differential in either direction, i.e., favoring the cases or the controls. It in some circumstances the case may not remember as well as controls because they are ill. It is also true that nondifferential exposure misclassification can, and probably does, occur in case-control studies. This is simply because of inability of the subjects to recall past life events that is not biased by disease status. In any case, the paper by Raphael (1987) does not demonstrate that "misclassification would be differential with high exposures reported by participants with disease" it only indicated that might occur and how to assess and control for this. The Panel should provide some direct evidence that case response bias occurs in epidemiologic studies of pesticides rather than just assuming that it does occur. You might want to look at Blair and Zahm in Epidemiology 1993:4:55-62 for a methodologic study that showed it did not occur. *EFSA Response:* *Same text as comment #212.* • 3107-3108. Might want to refer to the papers by Thomas and Coble, which are already cited in the document. They provide a validation of questionnaires used in the Agricultural Health Study. Since these questionnaires are much like those used in other studies, this validation suggests that this questionnaire approach works pretty well for farmers. *EFSA Response:* *Did this and found the following. "Reliability decreased as more detailed reporting was requested, i.e., for frequencies and amounts of substances used. Exact agreement for years and days per year of use of specific pesticides ranged from 50 to 75% " Blair A, Sandler D, Thomas K, et al. Disease and Injury Among Participants in the Agricultural Health Study. Journal of agricultural safety and health. 2005;11(2):141-150.*

				• 3124-3127: It is true that case-control and cross-section studies do not provide a temporal relationship that is directly available as in cohort studies. However, there are many things that can be done and are done in case-control studies to provide evidence this issue. This is another example of the Panel delivering a sweeping conclusion based on literally on a sound bite rather than evaluation of the literature. *EFSA Response:* *The quality of all studies irrespective of their design is weighed in the Weight of Evidence approach used and recommended by EFSA. Having said this some designs have proven more useful than others in the past and we will modify the text if the wording implies more than this.* • 3154-3156: Does the literature demonstrate that all of the multiple agents listed here cause all of the outcomes evaluated in the pesticide studies. If they do not, then they cannot confound. It is only essential to account for agents that cause the disease of interest. Another sweeping, and incorrect, conclusion from the Panel about epidemiology. *EFSA Response:* *Same text as comment #212.* • 3170-3172: With all of the limitations in epidemiology identified by the Panel in this document, it should made it clear why they believe that the associations between pesticides and Parkinson's disease and childhood leukemia are consistent and believable. What is different about the epidemiologic studies of these outcomes and evaluations from other pesticides and other diseases? *EFSA Response:* *The distinction between a proven association and a causal association is critical in interpreting epidemiological findings. The PPR Ppanel's publication on Parkinson's disease and childhood leukaemia was able to show the usefulness of establishing Adverse Outcome Pathways (AOPs) if one is able to provide a good AOP then the association is more plausibly a causal association. It does not prove causality but does make it more likely.* • 3223-3227: There are a number of methodologic studies that have evaluated the reliability and/or validity of the various techniques used to characterize human exposure to pesticides. These need to be reviewed and discussed by the Panel. • 3229-3235: Studies of non-malignant respiratory disease should be reviewed. There have also been many studies that evaluated breast cancer in relation to pesticides, mainly DDT. *EFSA Response:* *The use of DDT is banned in the EU and the present opinion is aimed at developing RA to improve assessment and reassessment of pesticides potentially coming to market or already on the market.* • 3238-3240: A hierarchy of study relevance is a not a sound approach. The hierarchy should first be based on the

				quality of determination of disease, assessment of exposure, study size, control for possible confounding and other biases. These factors are more important that the basic study design. There are many case-control studies that are superior to some cohort studies because of the manner in which they handle these important issues. *EFSA Response:* *Same text as comment #212.* • 3317-3320: The EPSA review can hardly be considered "comprehensive" since it was limited to publications only over a five year period. I have personally never heard of a "review" that restricted papers to such a narrow time window. A full explanation by the Panel is needed to explain this decision and to document that important data are not being excluded. I am sure important papers have been excluded. *EFSA Response:* *Same text as comment #212.* • 3365-3370: The Panel implies, but does not specifically state, that the second review by the Ontario College of Family Physicians found few positive associations. Such detail should be provided. The Panel should also characterize the differences between the first and second review that provide evidence for their conclusion that this is evidence of "over interpretation of epidemiological studies" in the first report. *EFSA Response:* *Same text as comment #212.*
214	Université de Bordeaux	FRA	9. Conclusions	• Exposure assessment is a major challenge in epidemiologic studies as the Panel points out. However, the conclusion is that this largely invalidates all epidemiologic data from integration into risk assessments. Most exposure misclassification is nondifferential (and for cohort studies almost entirely nondifferential). This means that if an association occurs between some pesticide or pesticide group, then the actual risk is larger than calculated in the study. This indicates that even studies with exposure misclassification can be used to provide guides to level of risk in humans. Although differential exposure misclassification can occur in case-control studies, the potential direction for the bias in not unidirectional (it can artificially increase or decrease the estimates of relative risk) and one should not assume that direction misclassification always occurs. In fact, there are relative few documented examples of such bias. *EFSA Response:* *One of the reasons why many pesticide epidemiology studies are unhelpful for risk assessment is the lack of reliability in exposure assessment. Risk assessment needs data on a single pesticide, but there are studies addressing this point by "ever versus never" use of pesticides or by self-reporting the use of pesticides. In these cases, it is hard to know whether exposure was only to a single pesticide or to different pesticides (or any other chemical) as well as the intensity, duration and timing of use can vary from one study to another. This largely limits their validity for risk assessment because of the uncertainties in the exposure assessment, which is key step in the process or risk assessment. The context of the Scientific Opinion is the usefulness of epidemiology studies for risk*

				assessment, nor the scientific value of these studies, which is indirectly considered. • 3405-3414: This paragraph makes a sweeping conclusion regarding the literature on reliability and validity of pesticide exposure techniques other than human biologic monitoring, but the Panel provides no documentation of the findings from studies available that support their conclusion. *EFSA Response:* *The commenter is invited to carefully read the EFSA External Scientific Report on human biomonitoring, which was briefly summarized in Annex B (Bevan R, Brown T, Matthies F, Sams C, Jones K, Hanlon J, La Vedrine M. Human biomonitoring data collection from occupational exposure to pesticides –Final Report, EFSA supporting publication 2017:EN-1185. 207 pp. Available at http://www.onlinelibrary.wiley.com/doi/10.2903/sp.efsa.2017.EN-1185/pdf)* • 3421-3423: I wonder why HBM for organochlorine pesticides were not considered. There are a large number of epidemiological studies that have used HBM for these chemicals. *EFSA Response:* *See comment above. Briefly, organochlorine pesticides are no longer approved for use in the EU.* • 3552-3555: There are many other epidemiological studies of pesticides that can and have provided useful information in addition to the AHS and Children's Environmental Health and Disease Prevention Research Centers. This statement seriously underplays the magnitude of the sources of the literature on pesticides. *EFSA Response:* *Same text as part 1) of comment #223.* *The PPR Panel agrees with this comment, but the text in lines 3551-3555 states that US-EPA put additional emphases on those two sources of information.* *The externally sourced systematic review was the starting point for this study. Without you providing specific references it is impossible to explain for what reason the studies you allude to were excluded. This opinion recommends greater use of HBM to improve exposure assessment.* • 3755-3757: Some early AHS findings are based on small numbers, but not all. It depends on the incidence of the disease or outcome. *EFSA Response:* *Thank you for this comment, which is considered correct. However, lines 3755-3757 do not deal with the AHS.* • Effect size magnification/inflation 3869-4211: This section deals with the "noise" associated with statistical estimates of risks. The smaller the study the greater the noise. Point estimates epidemiology are typically accompanied by calculating confidence intervals, which provide a range based on statistical probability in which the "true" risk estimate is likely to occur. This interval shows that the true risk may be either larger or smaller than the

				one calculated from the study data. The discussion by the Panel focuses overwhelmingly on the inflated effect. It would have been possible to present tables and graphs showing the results of size minimization from small studies. Failure by the Panel to do this creates a very distorted message, i.e., epidemiology is largely afflicted with "size magnification." Some discussion here about how reviewers deal with this problem is needed. As is it appears that epidemiologists are unaware of the problem and do nothing to deal with it. Independent replication and pooling and/or meta-analyses provide protection against biases toward the null or from size inflation from single studies. *EFSA Response:* *Same text as first part of comment #222.* • 4088-4091: The Panel notes that underpowered studies can lead to inflation, there are other biases that may be in the opposite direction. They mention non-differential misclassification, as an example. This is correct, but they do not say the non-differential misclassification is a additional factor that may bias studies toward the null. This can, and probably does, occur in large and small studies. Underpowered studies cam also bias estimates toward the null for the same reason inflation might occur, i.e., statistical variability. So in keeping score of biases in epidemiologic studies, underpowered ones can bias estimates in either direction (the confidence interval provides an indication of the size of such biases) and non-differential misclassification of exposure or outcome would bias the estimates toward the null, but not away from the null. This document gives the impression that there is a balancing of the biases up and down and this is just not the case. *EFSA Response:* *The PPR Panel agrees with this comment, which in fact is addressed in lines 908-927 of the Opinion. However, the PPR Panel put more emphasis on effect size magnification/inflation (Annex D) because of its real impact on risk assessment and regulatory decisions.* • 4093-4094: The Panel notes that underpowered studies can "commonly (but not always) produce predictable biases toward the null". Exactly the wording can be used for size magnification, but the Panel does not. *EFSA Response:* *Same text as comment #223.* • 4151-4153: This sentence demonstrates the faulty reasoning by the Panel regarding the underpowered study effect. "Specifically: any discovered associations from an underpowered study that are highlighted or focused upon on the basis of passing a statistical or other similar threshold are systematically biased away from the null. This just is not correct. Based strictly on size, they are just as likely to biased toward the null. *EFSA Response:* *Same text as comment #223.* • 4161-4176: This paragraph has several assumptions that are not correct. First, noise from underpowered studies